

Python for Data Analysis - 48 hours

Course Overview:

Python is a versatile and user-friendly language ideal for those new to programming. Its well-designed structure and syntax make it a powerful tool for various applications, including data analysis, which has surged in popularity recently. While one need not be an expert in Python to analyze data, grasping the language's fundamentals is essential before progressing to practical applications.

Our course is precisely structured to establish a robust foundation in Python, along with its pivotal packages for data analysis tasks. This program is intended for data analysts and users with minimal coding experience who seek to effectively manage data analysis projects.

The curriculum encompasses basic to intermediate Python programming coupled with key libraries essential for data analysis Like **Numpy, Matplotlib & Pandas**:

Moreover, the course will guide learners through using Python in a Big Data context, offering an introduction to critical platforms like Hadoop and Spark, and focusing on the principal library for such infrastructures – Pyspark.

What Participants Gain?

- Writing clean, structured Python code for practical data analysis
- Using NumPy for efficient numerical operations and arrays
- Manipulating and analysing data with pandas (Series, DataFrames, GroupBy)
- Creating clear visualizations with Matplotlib and Seaborn
- Working with multiple data formats and APIs (JSON, Excel, web resources)
- Applying Python in data and big data contexts (intro to PySpark)

Prerequisites

- Basic math skills
- Basic to Intermediate Python Skills are the benefit

Target Audience

- Entry-level Data Analysts or analysts looking to strengthen their Python skills
- BI professionals who want to extend their toolkit beyond SQL and BI tools
- Data practitioners seeking to move from manual analysis to automated workflows
- Graduates of BI / SQL courses aiming to deepen analytical capabilities with code
- Developers and technical professionals looking to build a practical foundation in Python for data analysis

Course Contents:

Part I: Basic Python:

- Python Fundamentals
 - Introduction to Python
 - Core elements of Python programming
 - Configuring the Python development environment
- Understanding Data Types in Python
 - Numbers: Integer, Float, and Complex types
 - Strings: Character processing in Python
 - Boolean Values: True/False evaluations
 - The None Type: Usage and implications
- Collections in Python
 - Lists: Ordered and mutable collection of items
 - Tuples: Immutable sequences of values
 - Dictionaries: Key-value pairs for efficient data lookup
 - Sets: Unordered collections for unique elements
- Control Flow Mechanisms
 - Conditional statements with if...else
 - Iteration with for...in loops
 - Pythonic data processing with list comprehensions
 - Loop control with while, continue, and break
- Interfacing with Textual Data
 - Gathering user input
 - String formatting for output

Part II: Intermediate Python:

- Functionality Enhancement with Functions
 - Crafting user-defined functions
 - Utilizing *args and **kwargs for variable arguments
 - Leveraging built-in function capabilities
 - Streamlining code with lambda expressions
- Debugging and Error Handling
 - Identifying and fixing code errors
 - Managing exceptions for enhanced code reliability
- Managing Text Files
 - Reading from and writing to text files
- Object-Oriented Programming (OOP) Concepts
 - Introduction to classes and objects in Python
 - OOP Concepts and patterns

- The Extensive Standard Library
 - Module importation techniques
 - Working with datetime for time-sensitive operations

Part III: Python Tools for Data Analysis:

- Delving into Data Analysis with pandas
 - Harnessing the power of NumPy for array operations and broadcasting
 - Creating compelling data visualizations with matplotlib
 - Introduction to matplotlib objects and functionalities
 - Diverse plotting techniques for data representation
 - Advanced visualization with Seaborn enhancements
 - Mastering pandas for sophisticated data handling
 - Exploring Series and Index structures
 - Manipulating DataFrames for complex data operations
 - Efficient data aggregation with GroupBy
 - Generating insightful visualizations within pandas
 - Practical applications and use cases
- Utility Tools for Data Analysts
 - Parsing patterns with regular expressions (re)
 - Working with JSON for web data interchange
 - Establishing connections with APIs and data resources